

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-222411

(43)Date of publication of application : 21.08.1998

(51)Int.Cl.

G06F 12/00

(21)Application number : 09-023013

(71)Applicant : NIPPON TELEGR & TELEPH CORP <NTT>

(22)Date of filing : 05.02.1997

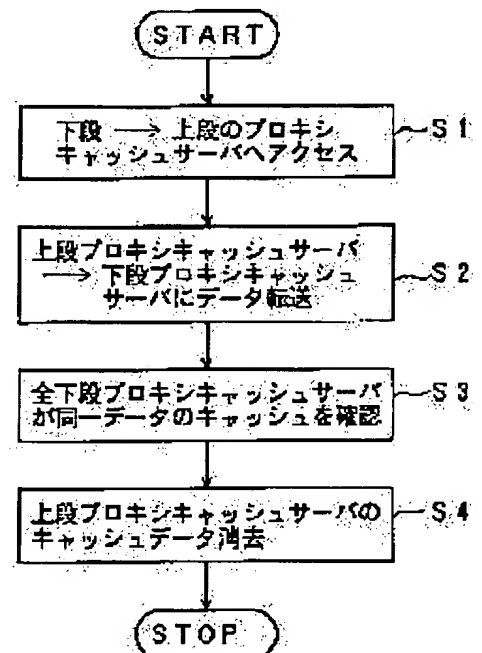
(72)Inventor : ASAKA TAKUYA

(54) PROXY CACHE SERVER CONTROL METHOD AND PROXY CACHE SERVER

(57)Abstract:

PROBLEM TO BE SOLVED: To leave data which is possibly referred to in an upper stage proxy cache server, to effectively use cache possible data capacity and to improve the hit ratio of cache data by erasing data of the upper stage proxy cache server when all lower stage proxy cache servers positioned immediately under the upper proxy cache server are recognized to cache same data.

SOLUTION: The lower stage proxy cache server connected to a user host accesses to the upper stage proxy cache server among the multistage proxy cache servers (step 1). Cache data is transferred from the upper stage cache server to the lower stage proxy cache server (step 2). When the upper stage proxy cache server recognizes that all the lower stage proxy cache servers cache same data (step 3), it erases cache data (step 4).



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-222411

(43) 公開日 平成10年(1998) 8月21日

(51) Int.Cl.⁸

G 0 6 F 12/00

識別記号

5 4 5

F I

G 0 6 F 12/00

5 4 5 Z

審査請求 未請求 請求項の数3 O L (全 6 頁)

(21) 出願番号 特願平9-23013

(22) 出願日 平成9年(1997) 2月5日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 朝香 卓也

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(74) 代理人 弁理士 伊東 忠彦

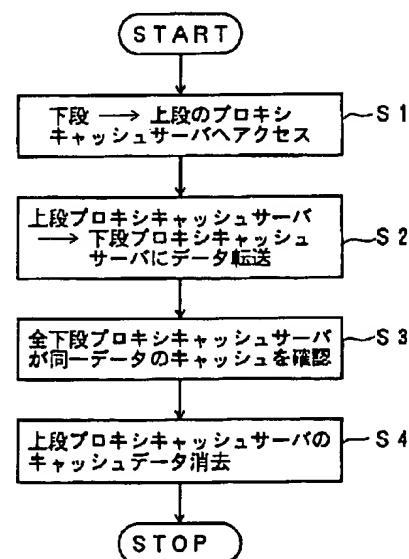
(54) 【発明の名称】 プロキシキャッシュサーバ制御方法及びプロキシキャッシュサーバ

(57) 【要約】

【課題】 上段プロキシキャッシュサーバには問い合わせがある可能性のあるデータだけが残し、キャッシュ可能データ容量を有効に利用することができ、結果として、キャッシュデータのヒット率を上げることが可能なプロキシキャッシュサーバ制御方法及びプロキシキャッシュサーバを提供することを目的とする。

【解決手段】 本発明は、多段プロキシキャッシュサーバのうち、上段プロキシキャッシュサーバにおいて、上段プロキシキャッシュサーバの直下に位置する全ての下段プロキシキャッシュサーバの全てが、同一のデータをキャッシュしたことが確認された時点で、上段のプロキシキャッシュサーバのデータを削除する。

本発明の原理を説明するための図



【特許請求の範囲】

【請求項1】 WWW等の蓄積型データ転送を目的とするネットワークの多段ブロキシキャッシュサーバネットワークにおいて、ブロキシキャッシュサーバを制御するブロキシキャッシュサーバ制御方法において、多段ブロキシキャッシュサーバのうち、上段ブロキシキャッシュサーバの直下に位置する全ての下段ブロキシキャッシュサーバが、ある同一のデータをキャッシュしたことが確認された時点で、上段のブロキシキャッシュサーバの該データを削除することを特徴とするブロキシキャッシュサーバ制御方法。

【請求項2】 前記多段ブロキシキャッシュサーバのうち、ユーザホストに接続される下段ブロキシキャッシュサーバから前記上段ブロキシキャッシュサーバにアクセスし、前記上段ブロキシキャッシュサーバから前記下段ブロキシキャッシュサーバに対してキャッシュデータを転送し、

前記上段ブロキシキャッシュサーバは、すべての下段ブロキシキャッシュサーバが同一のデータをキャッシュしたことを確認した時点で、該上段ブロキシキャッシュサーバのキャッシュデータを消去する請求項1記載のブロキシキャッシュサーバ制御方法。

【請求項3】 WWW等の蓄積型データ転送を目的とするネットワーク上の多段ブロキシキャッシュサーバから構成され、該ブロキシキャッシュサーバを制御するブロキシキャッシュサーバであって、

自ブロキシキャッシュサーバより下段のブロキシキャッシュサーバからのアクセスに基づいてキャッシュデータを転送するデータ転送手段と、

自ブロキシキャッシュサーバより下段のブロキシキャッシュサーバの全てが、前記キャッシュデータと同一のキャッシュデータをキャッシュしたか否かを判定するキャッシュ判定手段と、

前記キャッシュ判定手段において、キャッシュされたことが判定されると、上段のブロキシキャッシュサーバのデータを削除するキャッシュデータ削除手段とを有することを特徴とするブロキシキャッシュサーバ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ブロキシキャッシュサーバ制御方法及びブロキシキャッシュサーバに係り、特に、WWW等の蓄積型データ転送を目的とするネットワークにおいて、ブロキシキャッシュサーバに蓄積されるキャッシュデータを効率よく蓄積するための多段に接続されたブロキシキャッシュサーバの制御方法及びブロキシキャッシュサーバに関する。

【0002】

【従来の技術】 各ユーザは、ブロキシキャッシュサーバを経由し、目的とするデータを持つWWWページ等のデ

ータを配備しているデータ蓄積サーバにアクセスする。あるユーザがあるデータ蓄積サーバにアクセスし、データ蓄積サーバからユーザ側へデータ転送を行うときに、ブロキシキャッシュサーバは、当該データをキャッシュする。他のユーザが当該データを要求するときに、ブロキシキャッシュサーバは、データ蓄積サーバのデータの代わりに、キャッシュされたデータをユーザに転送する。このブロキシキャッシュサーバが図6に示すように、多段接続されているなら、上段のブロキシキャッシュサーバは、複数の下段ブロキシキャッシュサーバがキャッシュしているデータ全てをキャッシュすることができる。あるユーザがあるデータを取得しようとした時、下段ブロキシキャッシュサーバがキャッシュしていなかったとしても、上段ブロキシキャッシュサーバがデータをキャッシュしていれば、当該ユーザは、上段ブロキシキャッシュサーバからデータを取得することができる。結果として、キャッシュデータのヒット率を上げることができる。

【0003】 従来のツリー上に多段接続されたブロキシキャッシュサーバでは、各ブロキシキャッシュサーバは、独立にデータのキャッシュ及びキャッシュ消去を行っている図6の構成における上段及び下段のブロキシキャッシュサーバがキャッシュしていなかったら、あるユーザがデータ蓄積サーバのデータ取得時において、上段及び下段のブロキシキャッシュサーバは、共に当該データをキャッシュする。キャッシュデータを消去する場合には、各ブロキシキャッシュサーバは、「現在時刻-データのキャッシュ時刻」と閾値を比較して、閾値を越えていれば、消去する。あるいは、新しいデータをキャッシュするときに、ブロキシキャッシュサーバのキャッシュ可能データ容量を越えるなら、キャッシュデータ置き換えアルゴリズムに従い消去される。置き換えアルゴリズムとしては、例えば、キャッシュ時刻の古いものあるいは、データ量の大きなものを消去するといったものがある。

【0004】 以上をツリー上に多段接続されたブロキシキャッシュサーバの処理をまとめると以下のようになる。なお、従来のブロキシキャッシュサーバでは、上下段のブロキシキャッシュサーバともに同じ処理を行う。

40 〈データ蓄積時〉

・下段・上段ブロキシキャッシュサーバ共に当該データをキャッシュしていなかった場合：上段及び下段のブロキシキャッシュサーバとともにキャッシュすると同時に、当該データをユーザに転送する。

【0005】 ・上段ブロキシキャッシュサーバだけが当該データキャッシュしていた場合：下段ブロキシキャッシュサーバがキャッシュすると同時に、当該データをユーザに転送する。

・下段が当該データをキャッシュしていた場合：下段ブロキシキャッシュサーバが、当該データをユーザに転送

【0006】〈データ消去時〉

- ・キャッシュ可能データ容量を越えて新しくデータをキャッシュしようとする場合：キャッシュデータ置き換えアルゴリズムに従い消去する。

【発明が解決しようとする課題】しかしながら、上記従来のツリー上に多段接続されたブロキシキャッシュサーバでは、上段のブロキシキャッシュサーバのデータ消去時、特に、ブロキシキャッシュサーバのキャッシュ可能データ容量を越える場合に、キャッシュ時刻の古いもの、あるいは、データ量の大きなものを消去することは効率的でない。なぜなら、多段接続されたブロキシキャッシュサーバでは、下段ブロキシキャッシュサーバ全てが、キャッシュしているデータは、もはや、上段ブロキシキャッシュサーバに対する問い合わせがないにもかかわらず、上段ブロキシキャッシュサーバが当該データをキャッシュデータとして保持し続けることによる。

【0009】

【0010】また、本発明は、多段ブロキシキャッシュサーバのうち、ユーザホストに接続される下段ブロキシキャッシュサーバから上段ブロキシキャッシュサーバにアクセスし（ステップ1）、上段ブロキシキャッシュサーバから下段ブロキシキャッシュサーバに対してキャッシュデータを転送し（ステップ2）、上段ブロキシキャッシュサーバは、すべての下段ブロキシキャッシュサーバが同一のデータをキャッシュしたことを確認した時点で（ステップ3）、該上段ブロキシキャッシュサーバの

【 0 0 1 1 】 図 2 は、本発明の原理構成図である。本発明は、WWW等の蓄積型データ転送を目的とするネットワーク上の多段ブロキシキャッシュサーバ100、200から構成され、該ブロキシキャッシュサーバ100、200を制御するブロキシキャッシュサーバであって、自ブロキシキャッシュサーバ100より下段のブロキシキャッシュサーバ200からのアクセスに基づいてキャッシュデータを転送するデータ転送手段10と、自ブロキシキャッシュサーバ100より下段の全てのブロキシキャッシュサーバ200が、同一のキャッシュデータをキャッシュしたか否かを判定するキャッシュ判定手段20と、キャッシュ判定手段20において、キャッシュされたことが判定されると、自ブロキシキャッシュサーバ100のデータを削除するキャッシュデータ削除手段30とを有する。

【００１２】上記のように、本発明では、アクセスのあった下段ブロキシキャッシュサーバにキャッシュデータを転送し、当該キャッシュデータがキャッシュされたかを確認し、その確認がとれた時点で、転送元のブロキシキャッシュサーバのキャッシュデータを削除することにより、上段ブロキシキャッシュサーバに、アクセスされる可能性のあるデータのみが残されることになる。

【発明の実施の形態】図3は、本発明のプロキシキャッシュサーバの構成を示す。同図に示すネットワークの図は、WWWページ等の蓄積型データ転送を目的とするネットワークを表す。同図に示す多段プロキシキャッシュサーバは、上段プロキシキャッシュサーバ1、下段プロキシキャッシュサーバ2、3及び、WWWページの取得を要求するユーザのホスト6から構成される。

【0014】上段ブロキシキャッシュサーバ1は、キャッシュデータ管理部4と下段ブロキシキャッシュサーバ情報管理部5を有する。キャッシュデータ管理部4は、WWWサーバのWWWページのデータを蓄積し、管理する。下段ブロキシキャッシュサーバ情報管理部5は、どのキャッシュデータをどの下段ブロキシキャッシュサーバが保持しているかという情報を管理する。

【0015】上段ブロキシキャッシュサーバ1に対して下段ブロキシキャッシュサーバA2からキャッシュデータの要求があると、上段ブロキシキャッシュサーバ1のキャッシュデータ管理部4は、蓄積しているキャッシュデータを下段ブロキシキャッシュサーバA2に転送すると共に、キャッシュデータを下段ブロキシキャッシュサーバA2に転送したことを下段ブロキシキャッシュサーバ情報管理部5に通知する。これにより、下段ブロキシキャッシュサーバ情報管理部5は、キャッシュデータの転送先を管理する。

【0016】なお、上段ブロキシキャッシュサーバ1、
下段ブロキシキャッシュサーバA2、下段ブロキシキャ

ッシュサーバB3共に、従来のキャッシュ制御方法に加えて、本発明による上段ブロキシキャッシュサーバ1のキャッシュデータの消去も行う。図4は、本発明の上段ブロキシキャッシュサーバのキャッシュデータ消去動作のシーケンスチャートである。

【0017】ステップ101) ユーザホスト6は、下段ブロキシキャッシュサーバ2へアクセスする。

ステップ102) 下段ブロキシキャッシュサーバ2において望むWWWページのデータがなく、下段ブロキシキャッシュサーバ2は、上段ブロキシキャッシュサーバ1へアクセスする。

【0018】ステップ103) 上段ブロキシキャッシュサーバ1は、キャッシュデータを下段ブロキシキャッシュサーバ2に対して転送する。

ステップ104) 既に、下段ブロキシキャッシュサーバ2、3(図3の例では、下段ブロキシキャッシュサーバA、B)が同一のデータをキャッシュすることを上段ブロキシキャッシュサーバ1内の下段ブロキシキャッシュサーバ情報管理部5が検出したら、上段ブロキシキャッシュサーバ1内のキャッシュデータ管理部4にある当該キャッシュデータを消去する。

【0019】

【実施例】以下、本発明の実施例を前述の図3及び図4に基づいて説明する。上記の一連の動作を図3の例に適用すると、上段ブロキシキャッシュサーバ1には、当該ブロキシキャッシュサーバ1の下段に位置する下段ブロキシキャッシュサーバA2と下段ブロキシキャッシュサーバB3のみが接続されている。

【0020】ここで、ユーザホスト6からデータ要求された下段ブロキシキャッシュサーバA2から上段ブロキシキャッシュサーバ1に対してデータ要求を発行すると、上段ブロキシキャッシュサーバ1のキャッシュデータ管理部4は、下段ブロキシキャッシュサーバA2に対してキャッシュデータを転送すると共に、転送した下段ブロキシキャッシュサーバA2の情報(=A)及びキャッシュデータ識別子(ID=a)を、下段ブロキシキャッシュサーバ情報管理部5に通知する。これにより、下段ブロキシキャッシュサーバ情報管理部5は、転送したサーバ名「A」とキャッシュデータ識別子(ID=a)、及び時刻等を記録する。

【0021】次に、下段ブロキシキャッシュサーバB3から上段ブロキシキャッシュサーバ1に下段ブロキシキャッシュサーバA2から要求されたデータと同様のデータ要求が発行されると、上段ブロキシキャッシュサーバ1のキャッシュデータ管理部4は、当該データ要求に対応するキャッシュデータを下段ブロキシキャッシュサーバB3に転送する。そして、下段ブロキシキャッシュサーバ情報管理部5に対して、データ転送した下段ブロキシキャッシュサーバB3の情報(=B)及びキャッシュデータ識別子(ID=a)を下段ブロキシキャッシュサ

ーバ情報管理部5に通知する。これにより、下段ブロキシキャッシュサーバ情報管理部5は、転送したサーバ名「B」とキャッシュデータ識別子(ID=a)、及び時刻等を記録する。

【0022】上段ブロキシキャッシュサーバ1に接続されている下段ブロキシキャッシュサーバは、下段ブロキシキャッシュサーバA2と下段ブロキシキャッシュサーバB3のみであるため、下段ブロキシキャッシュサーバ情報管理部5は、すべての下段ブロキシキャッシュサーバA2、下段ブロキシキャッシュサーバB3に対して同一のキャッシュデータ(ID=a)が転送されたことを確認する。これにより、キャッシュデータ管理部4に格納されている上記キャッシュデータ識別子(ID=a)のキャッシュデータを消去する。

【0023】なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

【0024】

【発明の効果】以下に、本発明の効果を説明する。まず、図6のモデルを用いて説明する。同図において、上段ブロキシキャッシュサーバを1台、下段ブロキシキャッシュサーバを2台、総ユーザ数を200のもとで、各下段ブロキシキャッシュサーバにそれぞれ100ユーザずつがアクセスするものとする。各ユーザからのデータへのアクセスは、ポアソン到着するものとし、その平均アクセス時間間隔を100とした。ここでは、データiへのアクセスは、

$$\lambda(i) = q(1-q)^{i-1} \quad (1)$$

に従い発生するものとした。但し、qは、アクセス先データのインデックスiの平均を決定するパラメータで、その平均値は1/qによって得られる。ここでは、q=0.99とした。これにより、iが小さいほどアクセス率が高いことになる。また、図6でのユーザは、データ蓄積サーバは持っていないものとする。500単位時間でタイムアウトし、タイムアウトしたキャッシュデータは消去される。また、ブロキシキャッシュサーバ内のキャッシュデータ数は、ブロキシキャッシュサーバのエントリ数を上限とし、データ量による上限は考えない。また、エントリ数分のキャッシュデータが既にある状態で、新しいデータがキャッシュされる場合には、従来方法では、置き換えアルゴリズムLRU(Least Recently Used)に従い、古いデータが消去される。置き換えアルゴリズムLRUとは、できるだけ最終アクセス時刻が古いキャッシュデータから消去される。また、上段・下段ともブロキシキャッシュサーバのエントリ数上限は同一とした。

【0025】図5は、本発明の効果を示すための評価結果である。各ブロキシキャッシュサーバがキャッシュできるエントリ数が変化した場合のエントリあたりのヒット率を求めた。同図の評価結果からわかるように、上段

プロキシキャッシュサーバのヒット率について、提案方法は、従来方法に比べて有効であることが判る。

【0026】本発明の方法による上段プロキシキャッシュサーバのキャッシュデータ消去手順は、従来からのキャッシュ制御方法と併用することが可能である。その理由は、本発明の方法における消去実行時点が従来の消去実行時点と異なることによる。また、本発明では、上段プロキシキャッシュサーバだけの機能として実現できる。上段プロキシキャッシュサーバは下段プロキシキャッシュサーバのキャッシュするデータを知ることが常にできることから、本発明では、上段プロキシキャッシュサーバと下段プロキシキャッシュサーバ間で特別なプロトコルは必要としない。

【図面の簡単な説明】

【図1】本発明の原理を説明するための図である。

【図2】本発明の原理構成図である。

【図3】本発明のプロキシキャッシュサーバの構成図である。

*【図4】本発明の上段プロキシキャッシュサーバのキャッシュデータ消去動作のシーケンスチャートである。

【図5】本発明の効果を示すための評価結果である。

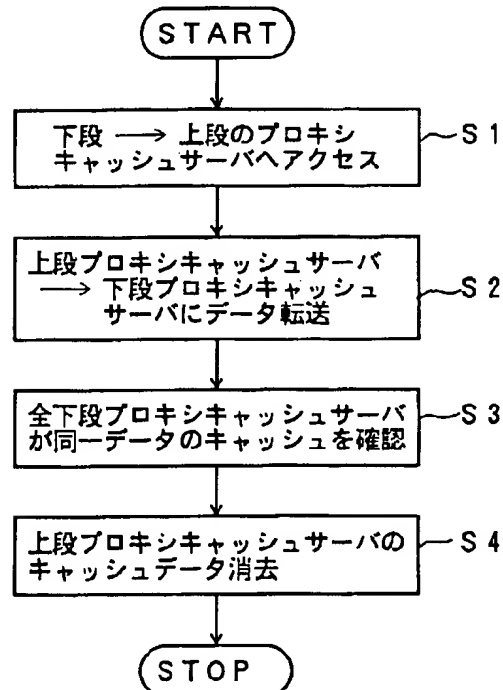
【図6】多段接続されているプロキシキャッシュサーバの例である。

【符号の説明】

- 1 上段プロキシキャッシュサーバ
- 2 下段プロキシキャッシュサーバA
- 3 下段プロキシキャッシュサーバB
- 4 キャッシュデータ管理部
- 5 下段プロキシキャッシュサーバ情報管理部
- 6 ユーザホスト
- 10 データ転送手段
- 20 キャッシュ判定手段
- 30 キャッシュデータ削除手段
- 100 上段プロキシキャッシュサーバ
- 200 下段プロキシキャッシュサーバ

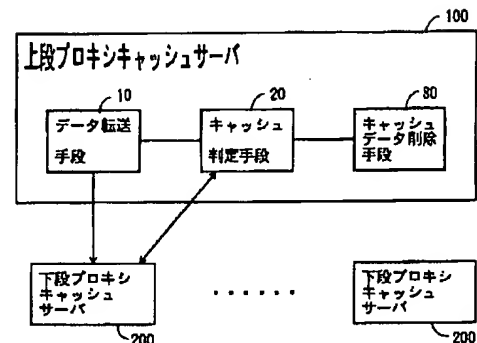
【図1】

本発明の原理を説明するための図



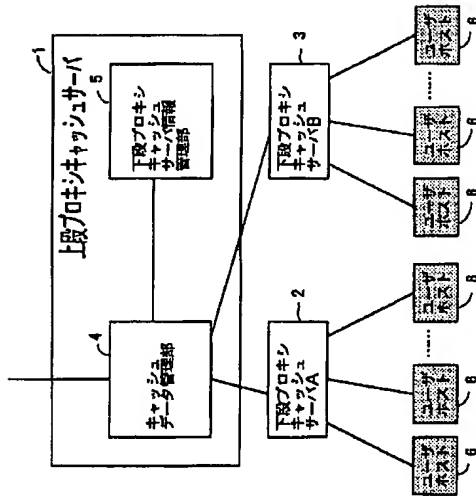
【図2】

本発明の原理構成図



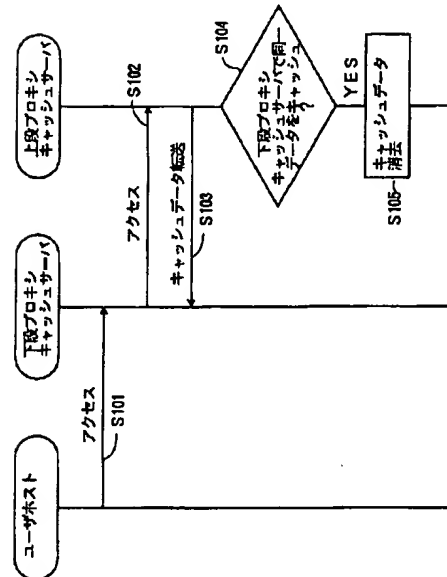
【図3】

本発明のプロキシキャッシュサーバの構成図



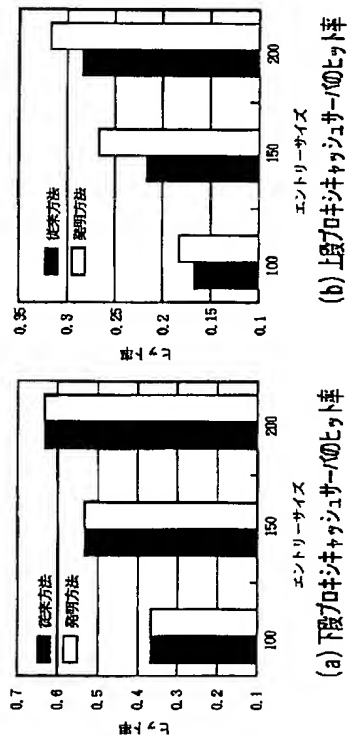
【図4】

本発明の上段プロキシキャッシュサーバの
キャッシュデータ消去動作のシーケンスチャート



【図5】

本発明の効果をj示すためのj平均結果



【図6】

多段接続されているプロキシキャッシュサーバの例

